

**编者按:**古希腊哲学家赫拉克利特的名言“人不能两次踏进同一条河流”,意谓万事万物皆处于永恒的运动变化当中。出自《吕氏春秋》的成语“流水不腐,户枢不蠹”,则是说事物只有处于运动变化当中,才会有充沛的活力,其生命力才能经久不衰。事物如此,对事物的认识和探究同样如此。一项研究、一门学问,如果因循守旧、陈陈相因是不可能创造新知的,也是不可能富有生机与活力的。相反,抱着矢志创新的态度,与时俱进地观察新现象、变换新角度、引入新方法,治学境界往往豁然开朗、面目一新。学术的生命在于创新,相信本专题中的三篇文章在这方面能够给人一些启迪。

## 机器人写作:AI 应用的技术反思

李红秀

**【摘要】**在大数据时代,人工智能(Artificial Intelligence,简称“AI”)技术正在成为社会生存的基础,机器人写作在新闻界的应用越来越广泛。文章研究机器人写作的发展现状和 AI 技术中的自然语言生成(Natural Language Generation,简称“NLG”)系统原理,探讨机器人写作与记者写作之间的关系,分析机器人生成新闻的可信度问题,指出社会物理学时代机器人写作面临的困境,进而指出,对 AI 技术的应用要持理性态度,对机器人写作的使用要有一定的限度。

**【关键词】**机器人写作;AI 技术;记者写作;可信度

伴随着计算机科学的快速发展,基于人工智能(Artificial Intelligence,简称“AI”)技术的各

种机器人不断被研发出来,机器人在社会生活中的使用越来越广泛。目前,机器人主要用于救

---

**作者简介:**李红秀,重庆交通大学人文学院教授,巴渝学者特聘教授,文学博士,硕士生导师,美国德瑞克大学(Drake University)访问学者。研究方向:新媒体传播和融媒体新闻。

灾、军事、工业、医学、教育等领域,比如搜索机器人、太空机器人、排毒机器人、手术机器人、学习机器人等。不仅如此,机器人还可以像人一样写作,而将写作机器人运用到新闻领域,在新闻界引起了轩然大波。

机器人写作最早出现在美国。2006年,美国汤姆森公司开始使用机器人撰写经济和金融方面的新闻稿件。<sup>[16]</sup>2007年,美国科技公司 Automated Insights 开发了一款名叫“WordSmith”的软件。这款软件可以自行编写一些简单事件的新闻,比如体育、财经类的新闻,雅虎、美联社的相当一部分新闻就是由 WordSmith 编写的。《纽约时报》《华盛顿邮报》《洛杉矶时报》《卫报》近年来都已经将机器人写作不同程度地用于新闻采编的多个环节。在中国,2015年9月10日,“腾讯财经”开发了写稿机器人“Dreamwriter”;<sup>[2]</sup>同年11月,新华社迎来一位新“同事”——机器人“快笔小新”。<sup>[16]</sup>2017年1月17日,南方都市报推出的南都机器人“小南”,一秒钟就完成了春运稿件的写作。<sup>[3]</sup>2017年4月24日,据新华社报道,全球首个机器人记者——中国智能机器人“佳佳”诞生,并越洋采访了美国著名科技观察家凯文·凯利。<sup>[4]</sup>

对新闻从业者而言,“狼来了”不是谎言,机器人写作对新闻业的冲击是显而易见的,但是,机器人写作对新闻业的发展是福还是祸,目前还无法估量。本文思考的问题是:机器人写作的基本规律是什么,机器人写作是否会取代人类写作,机器人写作是否真的完美无缺,在机器人写作时代人类记者有何存在的价值。

## 一、机器人写作的发展与 AI 技术原理

目前,学界对机器人写作进行现象描述和介绍的很多,对机器人写作进行定义的却很少。机器人写作在新闻领域应用比较频繁,因此大多数学者热衷于研究机器人新闻写作。最先对

机器人新闻写作进行研究的是美国学者佐伯芙(Shoshana Zuboff)。她在1981年发表的《计算机成为工作替代者》一文中这样描述:“不难想象,人类的记者在不久的将来将被软件取代。事实上,这种转变已经存在,编辑已由软件编辑替换。电脑编程故事将由机器人记者完成。数字化如何影响记者?多技能记者今后的处境怎样?他们将如何应对即将到来的令人不安的状况?什么是机器人新闻?为什么要使用算法新闻学?机器人生成什么样的文本?算法新闻会取代人类记者吗?自动文本要考虑哪些伦理问题?……这些都需要作出回答。”<sup>①</sup>佐伯芙对机器人写作在新闻领域的应用作了大胆预测和思考,特别是提出了与机器人写作相关的概念,比如“机器人新闻”(Robot Journalism),“机器人记者”(Robot Journalist),“算法新闻学”(Algorithmic Journalism),“自动化新闻”(Automated Journalism)。不过,她并没有对这些概念进行解释和说明。

机器人写作是伴随着 AI 技术的发展而发展的,其核心是云计算和大数据(Big Data)分析,即从庞杂的资讯中找出最可能受市场关注的那部分,将算法生成的内容(Algorithmic Generated Content,简称“AGC”)用人们能够接受的格式呈现出来。整个过程是自动化的。这是计算机基于算法的数据分析和自我学习的技术,体现了 AI 技术的“算法的中立性”(Neutrality of Algorithms)。这一中立性相当于新闻报道的客观性。

赖特(Reiter)指出:“‘机器人写作’又叫‘算法写作’(Algorithmic Writing),主要来源于自然语言生成(Natural Language Generation,简称‘NLG’)技术领域,这是基于非语言输入(Non-linguistic Input)的自动生成文本的过程。”<sup>[5]</sup>通常,NLG 系统必须能够执行如下标准化的任务:首先,选择哪些信息来表达内容;其次,组织可用信息并确定结构文本,确定哪些信息被放

①参见:刘建明.对智能新闻若干问题的释疑[J].新闻爱好者,2017(11):4.

置在任何句子中都能够准确表达意思;最后,创建要显示的表达式,同时语句要符合语法规则。只有通过这样的系统设置,机器人写出的文本才可能语法正确,表达清晰。

最初的 NLG 系统只能生成非常简单且没有太多变化的文本。经过科学家多年的努力,计算机语言识别技术得到快速发展,NLG 系统已经能够生成各种各样的文本。

如今,AI 技术快速发展,诞生了 Yseop、CBS(Columbia Broadcasting System)互动媒体公司等各种 AI 企业;同时,涌现了各种各样的机器人写作软件,比如,一款名为“幻想记者”(Fantasy Journalist)的软件,“能写出非常深刻和比较人性化的文章,这类文章很难分辨出是人写的还是机器写的”<sup>[6]</sup>。

奥贝莱特(Ombelet)指出,完成机器人自动化写作一般需要四个参与者(Actors)——软件程序员(Software Programmer)、数据源(Data Source)、编辑(Editor)和出版者(Publisher),其作用各不相同。“软件程序员是为了开发内容的创造性算法,数据源要确保能够为算法提供足够的原始数据,编辑的工作是选择数据源和监督自动化进程,出版者是为了给他们的读者发布所需要的内容”<sup>[7]</sup>。当然,在机器人写作的具体实践中,软件程序员、编辑和出版者可以是同一个参与者。比如,2014年3月17日,美国加州发生了一次4.7级的地震,《洛杉矶时报》(Los Angeles Times)的记者施文克(Ken Schwencke)在地震发生3分钟后,利用“地震机器人”(Quakebot)撰写了一条新闻。

据美国地质调查局报道,星期一早上,加利福尼亚韦斯特伍德5英里(8公里)处发生了一次浅层4.7级地震。地震发生在太平洋时间上午6:25,深度为5.0英里。美国地质调查局称,震中距加利福尼亚州贝弗利山庄6英里,距环球城7英里,距圣莫

尼卡7英里,距萨克拉门托市348英里。在过去的10天里,震中附近没有发生3.0级以上的地震。这一信息来自美国地质勘探局地震通知服务,这篇文章是由作者编写的一个算法创建的。<sup>[8]</sup>

AI 技术就像人类一样能够“思考”,只要创建一个自动化识别算法的软件,机器人就能根据用户的需要“写出”让用户感兴趣的个性化文本,比如谷歌、百度的网页排名和推特的趋势排名。以大数据为基础的机器人写作能够在几秒钟之内生成用户需要的文本,比如股票价格、体育统计、天气预报、财务报告等,能在短时间内满足用户的需要。

## 二、机器人写作与记者写作

记者写作新闻依赖于一个基本公式:标题、导语、主体、背景,主要元素包括5W+1H,即谁(Who)、何时(When)、何地(Where)、何事(What)、何因(Why)、何果(How)。对于具体的新闻事件而言,首先是要寻找新闻故事,找到读者感兴趣的故事点,对故事不能简单地平铺直叙;其次是要选择写作对象,这是新闻报道的核心,并且要写出生动的细节,这些细节还必须按正确的顺序排列,从而达到结构化(Structured)的效果;最后是根据现实中的故事一步步地展开写作。

随着机器人写作在新闻界越来越广泛的使用,一些学者开始对机器人新闻写作进行概念上的界定。学者喻国明指出:“所谓机器新闻写作是一种自然语言生成引擎,利用算法程序,通过采集大量的各种题材及高质量的数据,建立各种分类的庞大数据库,借助人工智能(AI)实现从数据到认识、见解和建议的提升和跨越,最后由机器自动生产新闻。”<sup>①</sup>申云则这样定义:“所谓机器新闻写作就是用编辑好的程序以及现有的新闻相关数据生成可视化的新闻,对新

①参见:申云.“机器人新闻写作”对新闻采编的机遇和挑战[J].今传媒,2016(11):115.

闻写作进行格局性改变。”<sup>[9]</sup>这两个定义虽然有所不同,但两位学者都充分认识到机器人新闻写作的核心要素是大数据和AI的应用,并且AI技术的渗透对于新闻传播生产方式和传播格局的改变是全方位、全过程的。

那么,随着机器人写作的到来,新闻界还需要人类记者吗?这在学术界引发了很大的争论。赖特认为,“这很大程度取决于经济利益,是雇佣和培养一个记者便宜呢,还是创造和维持一个机器软件便宜?经济决策通常在很大程度上取决于生成文本的数量”<sup>[9]</sup>。言下之意,机器人写作能否取代人类记者写作主要由市场决定,特别是传媒行业会从经济效率的角度进行权衡。瑞典学者克莱瓦尔(Clerwall)的观点比较明确,“机器人新闻可以说是免费资源,记者只需关注来自软件的信息是否合乎要求……如果仅仅是发现和获取低层次的信息资料,那么计算机新闻(Computational Journalism)比记者的工作更有实用价值。人类记者应该更多地关注新闻的验证、解释和传播”<sup>[10]</sup>。克莱瓦尔主张机器人写作和人类记者写作应该有不同的分工。

有些学者明确反对用机器人写作代替人类记者写作,认为机器人写作缺乏灵活性,不能为故事提供分析,如果让机器人撰写非常规的故事,就很难写得准确生动。达伦(Dalen)指出:“机器人无法避免陈词滥调,缺乏创造性和幽默感,虽然有人可能不同意这种看法。”<sup>[11]</sup>达伦并不反对使用机器人写作,认为机器人可以从事日常生活中一些简单化和自动化的写作,但是深度报道只能依靠人类记者完成,因为当前的NLG技术还无法达到要求。

### 三、机器人写作的可信度

客观性、平衡性和中立性已经成为现代新闻的标准,而可信度(Credibility)是评估新闻报道的核心。新闻报道的可信度通常包括来源的可信度、信息的可信度和媒介的可信度三个方面。可以说,可信度是媒体信誉的关键因素。瓦

纳克尔(Vanacker)和贝尔马(Belmas)指出:“可信度是我们对新闻媒体的期望之一,指的是报道是准确无误的,即使要扩大概念的范围,也必须是如实报道。”<sup>[12]</sup>由此可见,可信度在新闻报道信息的重要性方面不言而喻,直接影响到受众对新闻媒体的信誉评价。

通常,受众对媒体的信任度比较高,因为记者报道的新闻信息具有毋庸置疑的可靠性。但是,如果媒体使用机器人写作,那么机器人写出的新闻作品具有可信度吗?正如前文所述,机器人写作是基于NLG系统自动生成文本,如果生成的文本在语法上和逻辑上有明显问题(这可能是由于NLG系统还有缺陷),这样的文本信息当然是不可靠的,普通人也容易辨识。问题是,如果机器人写作完成的文本没有明显的语法上的毛病,它的信息可以被认为是可靠的吗?

根据荷兰学者卡(Kaa)和克拉默(Krahmer)的实证研究可知,读者根本无法辨识专业记者撰写的新闻和机器人撰写的新闻,认为两类文章都是真实可信的。“通过对新闻来源可靠性的观察(计算机和记者),读者发现不了二者的不同,他们所感知的计算机作者撰写的文章和新闻记者撰写的文章都具有可靠性。一般来说,他们对计算机作者和新闻作者的专业能力都保持中立态度。”<sup>[13]</sup>克莱瓦尔同样做了实证研究。<sup>[14]</sup>他找来两篇文章,一篇是人类记者写的,另一篇是由软件生成的(Software-Generated)。他邀请44个人参加测试,其中27人为普通人,17人为专业记者。测试结果发现,无论是普通人还是专业记者,都无法准确判断哪篇是人类记者写的,哪篇是机器人写的。测试者普遍感觉:“软件生成的文本更具描述性,但却无聊,这些内容比较详实、客观、准确和可信。”换句话说,读者对机器人写出的文本的可靠性并不怀疑。

但是机器人写作的可信度问题并没有结束。普通人在无法区别机器人写出的文本和记者写出的文本的前提下,以类推的方式认可了

机器人写出的文本也具有可信性,那么,如果受众知道一个文本是机器人写的,另一个文本是人类记者写的,二者的可信度谁更高呢?美国学者舒尔茨(Schultz)和谢菲尔(Sheffer)用实证研究方法比较了报纸和机器人记者(Robot Reporter)之间的可信度。<sup>[13]</sup>研究发现,“受访者的负面情绪是一致的”。几位受访者表示:“人类记者已经因为错误报道而臭名昭著(特别是在推测、事实分析等方面),现在机器人写作会变得更糟。”另一个更直言不讳:“我永远不会。如果你买报纸或读文章,如果它是由一个畸形机器人(Freaking Robot)写的。你在开玩笑吧?那么多有才华的作家只写些鸡毛蒜皮的事,因为他们用机器人写作。机器人知道什么是修辞学?机器人知道什么是新闻中的人性元素吗?”研究结果令人惊讶。越来越多的新闻媒体开始使用机器人写作,却不知道大众对机器人写作的厌恶。

#### 四、社会物理学时代机器人写作的困惑

我们生活在以数据库为基础的社会系统极其复杂的时代,数据库不断增长,大数据已经成为人类社会必不可少的组成部分,因此,拉塔(Latar)称当前社会为“社会物理学时代”(Age of Social Physics)。彭特兰(Pentland)这样解释,“各种新科学不断涌现,一切可以自动化分析, AI 算法普遍使用,数十亿个微型社交活动(Micro Social Engagements)不断地通过我们的移动设备与在线平台(Online Platform)联系在一起,这些微活动(Micro Engagements)已经被数字化,并存储在无限的数据库(Data Silos)中”<sup>[14]</sup>。AI 技术是社会物理学时代的标志之一,越来越多的媒体、企业、组织机构使用机器人写作,因为 AI 算法可以利用数据库来将数据转化为故事。

不少人感叹,人类记者会被机器人取代,新闻界对人类记者的需求将越来越少。然而,任何新技术都是一把双刃剑,机器人写作绝不是万

能的,其缺陷和问题不容回避。

第一, AI 算法本身存在一定的局限性。AI 算法实际上是数据挖掘基础上的统计学分析。机器人写作的基础是 AI 算法,通过设置变量使本身并无意义的数据之间建立联系,生成文本。由机器人生成的文本有的有意义,有的没有意义和实际价值,有的甚至可能是错误信息。所以,由机器人写出的文本特别是新闻作品,必须通过测试和校验,验证过程最好由人类分析家或记者来做。当然,人类分析家或记者要学会激活可用的新验证工具(数据分析)。

第二, AI 技术缺乏人类的理解能力。人类的语言丰富多彩,尤其是人类的思考和语言中的隐喻、相互交流中的幽默等,这些都不是机器人能够完全理解的。语言在文化和社会环境中也随着时间的变化而变化,机器人能够自由地跟随变化吗?维诺格拉德(Vinograd)和弗洛雷斯(Flores)认为,“AI 理解人类语言的能力不可能超越‘官僚层面(Bureaucratic Level)’”<sup>[15]</sup>。特劳桑·马图(Trausan-Matu)解释, AI “像官僚一样没有同情心,它严格遵循机器规则”<sup>[16]</sup>。也就是说,一个机器人记者无法写出内容丰富且有深度的故事,而这恰好是人类记者的优势所在。

第三,机器人写作不具有创新性。无论是产品还是技术,都是人类的发明创造。发明需要的创新能力为人类所独有,正是人类的各种技术发明才使社会不断进步。AI 技术是人类发明的成果之一,但它本身不具有创新能力。机器人自己不能提问,自己不能思考问题,它所有的“问题”和“答案”都是人类软件编写者设置的算法指令。人类的创新活动是一个复杂的思维过程,前提是以问题为导向,从提出问题到分析问题直至解决问题,有时需要很长的时间。机器人写作可能提供新的知识(这一点也需要验证),但不能制定政策,也不能提出促使社会进步的有效建议,至少在短时间内无法用 AI 技术来实现。

第四,机器人写作无法解决作品的法律责任问题。首先,机器人写出的作品具有知识产权

吗?西方国家非常重视知识产权的保护,特别注重保护人类的写作成果,如果其出版发行,创作者可以获得经济收益。那么机器人写出的作品如果出版,受益人是谁?是软件公司还是使用者?或者是机器人“自己”?另外,机器人写作如果涉及法律纠纷,由谁承担责任?新闻记者一般会对其作品的真实性及法律纠纷承担责任,因为作品上有他们的署名。可是,机器人写出的作品如果出现错误,或者暴露他人隐私,甚至无端诽谤他人,由此引起的法律问题该由谁来承担后果?是软件公司,或是使用者,还是编辑人员?由于无法简单划分责任,受害人应该找谁索赔也是难以界定的问题。

### 五、机器人写作的限度思考

AI 技术不断发展并被越来越广泛地应用于人类社会的各个领域,我们的世界仿佛一夜之间变得越来越方便和快捷。但是,一些思想家和学者却对 AI 技术持谨慎的乐观态度。科罗德尼(Kolodny)提出了一个反乌托邦式的预言(Dystopian Prediction):“科幻小说长久以来的主题是,人类世界被机器人管理,包括一个‘反常’(Singularity)观念:计算机最终超越了人类的知识,并控制人类。对于世界上一些最优

秀的思想家来说,这种想法正在从幻想(Fantasy)变得更具‘似真性’(Plausibility)。包括著名科学家史蒂芬·霍金(Stephen Hawking)也提出了一个很直接的问题:我们有必要担心将来机器人会比我们所有人类都聪明,这不只是科幻恐怖电影的想象。”<sup>[17]</sup>如果滥用 AI 技术,这对人类世界不是福音而是灾难。因此,我们要防止人类用自己发明的技术毁灭人类世界。

同样,要防止基于 AI 技术的机器人写作普遍化、泛滥化。新闻是艺术和科学的结合体。“新闻工作的艺术性表现在寻找新的创意,报道一个有创新角度的新闻,反映新的思想、解决问题的新方法以及丰富生活的新途径。新闻工作的科学部分是运用分析工具来支持和验证提出的想法,这些想法来源于人类活动被记录和被存储的数据库。”<sup>[14]</sup>那么,机器人写出的新闻具有艺术性吗?答案当然是否定的。机器人写出的文章不具有独创性以及情感,有时甚至会因观点偏颇而引起法律纠纷。如果新闻媒体都依靠 AI 技术,新闻作品都依靠机器人写作,那么,正如海德格尔(Heidegger)所担心的那样,人类还能“诗意地栖居”在世界上吗?因此,我们对 AI 技术的应用要持理性态度,机器人写作的使用要有一定的限度。

#### 【参考文献】

- [1] 孙振虎,张馨亚.机器人新闻的发展与反思[J].电视研究,2016(06).
- [2] 百度百科.腾讯写作机器人[EB/OL].<https://baike.baidu.com/item/%E8%85%BE%E8%AE%AF%E5%86%99%E4%BD%9C%E6%9C%BA%E5%99%A8%E4%BA%BA/18642443>.
- [3] 白龙.机器写作的新突破及思考——以南方都市报写作机器人“小南”为例[J].青年记者,2017(10月中):94.
- [4] 新华社.全球首次出现机器人记者 对战到对话 AI 渐入佳境[EB/OL].(2017-04-25).<http://sports.sohu.com/20170425/n490648196.shtml>.
- [5] Ehud Reiter, Robert Dale. Building Applied Natural Language Generation Systems [J]. Natural Language Engineering, 1997(01):57-87.
- [6] Hille A. J. van der Kaa, Emiel Kraemer. Journalist versus News Consumer: The Perceived Credibility of Machine Writ-

- ten News[C]. Proceedings of the Computation + Journalism conference, 2014.
- [7] Pieter-Jan Ombelet, Aleksandra Kuczerawy, Peggy Valcke. Employing Robot Journalists: Legal Implications, Considerations and Recommendations[C]. WWW 2016 Companion, April 11–15, 2016, Montréal, Québec, Canada.
- [8] Paul Field. Artificial Intelligence—Ten Things You Never Knew [EB/OL]. (2016–10–5). <https://www.linkedin.com/pulse/artificial-intelligence-ten-things-you-never-knew-paul-field>.
- [9] 申云.“机器人新闻写作”对新闻采编的机遇和挑战[J]. 今传媒, 2016(11): 115.
- [10] Christer Clerwall. Enter the Robot Journalist[J]. Journalism Practice, 2014(05): 519–531.
- [11] Arjen Van Dalen. The Algorithms behind the Headlines[J]. Journalism Practice, 2012(05–06): 648–658.
- [12] Bastiaan Vanacker, Genelle Belmas. Trust and the Economics of News[J]. Journal of Mass Media Ethics, 2009(02–03): 110–126.
- [13] Brad Schultz, Mary Lou Sheffer. Newspaper Trust and Credibility in the Age of Robot Reporters [J]. Journal of Applied Journalism & Media Studies, 2017(02): 339–355.
- [14] Noam Lemelshtrich Latar. The Robot Journalist in The Age of Social Physics :The End of Human Journalism?[M]//Gali Einav, Ed. The New World of Transitioned Media: Digital Realignment and Industry Transformation. Switzerland: Springer International Publishing, 2015: 65–80.
- [15] Terry Vinograd, Fernando Flores. Understanding Computers and Cognition: a New Foundation for Design Norwood[M]. NJ: Addison–Wesley, 1987.
- [16] Stefan Trausan–Matu. Human Language and the Limits of Artificial Intelligence: A New Religion–Science Relation[J]. Science and Religion: Global Perspectives, 2015(June): 4–8.
- [17] C. Kolodny. Stephen Hawking is Terrified of Artificial Intelligence [EB/OL]. (2014–05–05). [http://www.hufngtonpost.com/2014/05/05/stephen-hawking-artificialintelligence\\_n\\_5267481.html](http://www.hufngtonpost.com/2014/05/05/stephen-hawking-artificialintelligence_n_5267481.html).